X. Appendix C

From the principle of heredity to the molecular elucidation of inborn errors of metabolism

A. Heredity

Aristotle: heredity is based on a transfer of information.

The capacity of plants and animals to pass on their own qualities to their offspring is so obvious that it must rank among the earliest human observations (Fig. X-1a). This observation might also be one of the most important steps in enabling the transition from nomadic, food-gathering groups of people to sedentary, agricultural-urban societies around 8000 BC.

Aristotle (384 – 322 BC, Fig. X-1b) was the first to recognise that biological inheritance could not be the passage of body samples through the generations, as was believed earlier by Hippocrates (460 BC – 380 BC), but instead must involve the transfer of information for the embryonic development of the individual. It took 23 centuries before Gregor Mendel would discover an experimental basis for this fundamental principle.

B. Classical genetics

The era of classical genetics started with Mendel's introduction of statistical analysis in crossing experiments (1866), and ended with Beadle and Tatum's introduction of biochemical analysis in genetics (1941).

B.1. Gregor Mendel and the garden pea

G. Mendel: hereditary traits are passed on as discrete units.

In 1866, Gregor J. Mendel (1822 – 1884, Fig. X-1c) published his paper 'Experiments on plant hybrids'[1354]. Using the garden pea, *P. sativum*, he was the first to statistically evaluate qualitative traits. From his results, he deduced that the hereditary traits of the pea are carried and passed on to the progeny as discrete units. Each pea must possess a homologous pair of such units, one received from the pollen, one from the ovum. He also introduced the con-

cepts of dominant and recessive inheritance. Ironically, Charles R. Darwin (1809 – 1882), a contemporary of Mendel, was never aware of Mendel's 'hereditary units', on which, as we now know, his theory of natural selection operates.

With the rediscovery of Mendel's laws of inheritance in 1900 by Hugo M. de Vries (1848 – 1935), Erich von Tschermak-Seysenegg (1871 – 1962) and Carl E. Correns (1864 – 1933)[1355], the study of heredity finally gained momentum. Due to the observation of the cell with its central nucleus, and the study of chromosomes and their behaviour in mitosis and meiosis (all ~1850s – ~1890s), these laws could finally be interpreted. Terms such as 'genetics', 'gene', 'allelomorph' (later 'allele'), 'zygote', 'homo-' and 'heterozygote', 'genome' and 'mutation' first appeared in this era.

B.2. Thomas Morgan and the fruit fly

T. Morgan: discovery of X and Y, linkage, recombination, genetic mapping and the concept of a fixed sequence of genes along a linear chromosome.

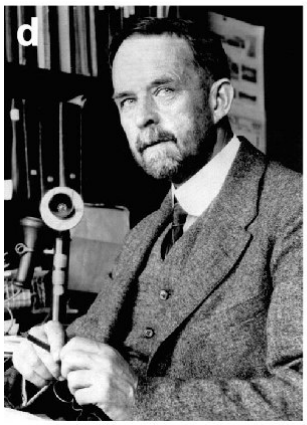
With the study of the fruit fly, *D. melanogaster*, started in 1910 by Thomas H. Morgan (1866 – 1945, Fig. X-1d), the next major steps in genetic research were taken. He discovered that male flies, as opposed to female flies, do not carry four pairs of homologous chromosomes. Rather, they have three such pairs and two chromosomes with a dissimilar morphology, one of which is akin to that constituting the fourth pair in female flies. The sex chromosomes (called X and Y) were discovered. Morgan recognised that sex is

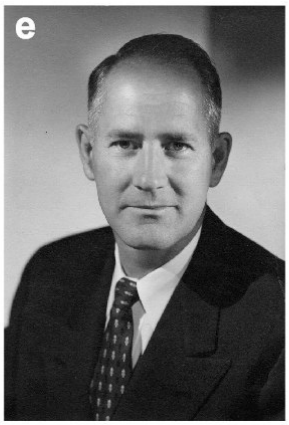
Fig. X-1. *Next page.* **A pictorial history of classical and molecular genetics.**

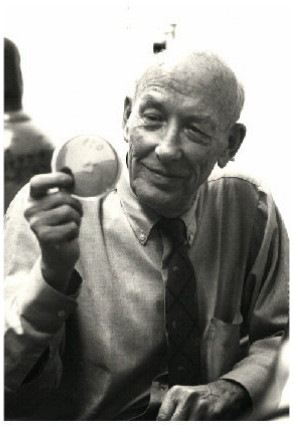
a. 'Heredity' by Thomas Hardy (1840 – 1928). First published in 'Moments of Vision and Miscellaneous Verses', Macmillan, 1917.
b. Aristotle (384 – 322 BC), oil on canvas, 1811, by Francesco Hayez (1791 – 1882).
c. Gregor Mendel (1822 – 1884).
d. Thomas Morgan (1866 – 1945).
e. George Beadle (1903 – 1989, left) and Edward Tatum (1909 – 1975, right).
f. Phoebus Levene (1869 – 1940).
g. Oswald Avery (1877 – 1955).
h. Francis Crick (1916 – 2004, left) and James Watson (°1928, right) walking along the Backs, Cambridge, UK.
i. Linus Pauling (1901 – 1994, left), Frederick Sanger photographed at the Wellcome Trust Sanger Institute, Cambridge, UK (°1918, middle) and Vernon Ingram (1924 – 2006, right).

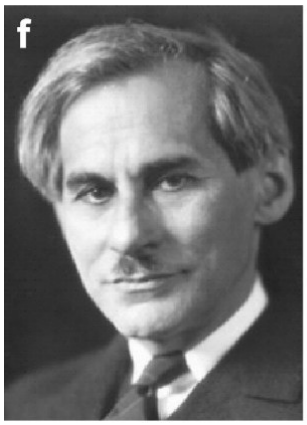
Images from Galleria dell'Accademia, Venice, Italy (b); http://www.biologie.uni-hamburg.de (c); University Archives, Columbiana Library, Columbia University, New York, USA (d); Special Collections, The Valley Library, Oregon State University, Corvallis, USA (e – left, g, h, i – left and i – right); Rockefeller Archive Center, Stanford University, Stanford, USA (e – right, f) and http://www.sanger.ac.uk (i – middle).

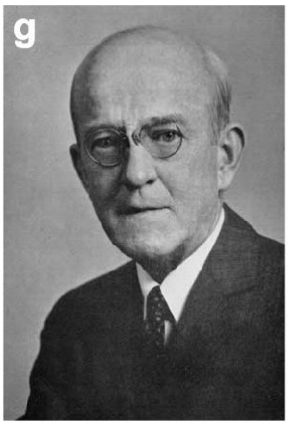
a *I am the family face:*
Flesh perishes, I live on,
Projecting trait and trace
Through time to times anon,
And leaping from place to place
Over oblivion.
The years-heired feature that can
In curve and voice and eye
Despise the human span
Of durance - that is I:
The eternal thing in man,
That heeds no call to die.

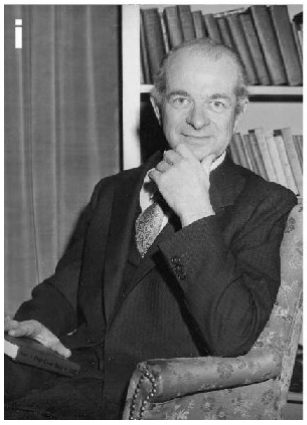


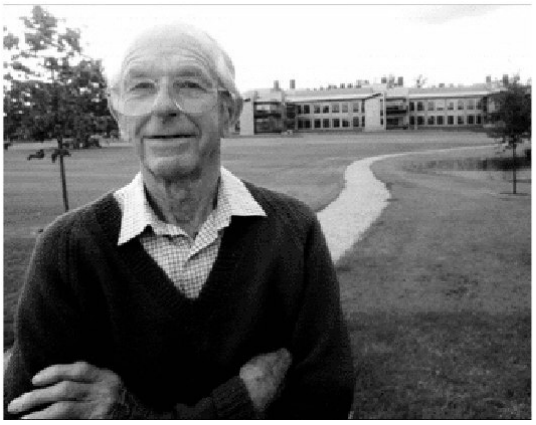


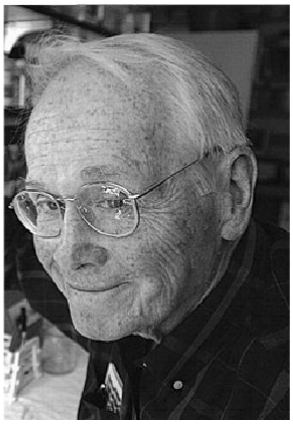












inherited as a simple Mendelian trait, with offspring always getting an X chromosome from their mothers and either an additional X from their fathers in case of female offspring or a Y in case of male offspring. Females are XX, males are XY, a finding that was soon confirmed in mammals.

By looking for mutations (first spontaneous and later X-ray-induced), Morgan established a collection of different mutants. One of them had white eyes, and through genetic crosses and statistical analysis, much like Mendel had done, he found that the trait segregated with the X chromosome. Eye colour was the first trait to be linked to a particular molecular entity, the X chromosome. Crossing flies with mutations on the same chromosome, he discovered 'linked genes' and recombination, which he – correctly – interpreted as crossing-over of homologous chromosomes, observed earlier during meiotic divisions of salamander sexual tissues. He then went on to use recombination frequencies between mutant loci as a measure for genetic distance and established the first genetic map in history[1356]. As a 'side-effect', Morgan's work implied that genes were distributed in a fixed sequence along a linear chromosome.

B.3. George Beadle, Edward Tatum and bread mould

G. Beadle and E. Tatum: one-gene-one-enzyme theory.

In the 1940s George W. Beadle (1903 – 1989, Fig. X-1e) and Edward L. Tatum (1909 – 1975, Fig. X-1e) combined genetic crosses with biochemical analysis in the bread mould *N. crassa*. Using selective media, they isolated large numbers of mutants defective in the synthesis of single biological components. Upon biochemical analysis, they found that these mutants are defective in a single step in the reaction sequence leading to the biological molecule. Genetic crosses revealed that these mutants carried mutations in single genes, which lead Beadle and Tatum to postulate their one-gene-one-enzyme theory[1357]. Not only were they the first to experimentally prove a relation between genes and proteins, but their theory also paved the way for the next era in genetic research, that of molecular genetics.

C. Molecular genetics

With the increased application of biochemical methods in genetics, cumulating in the nature and eventually the structure of the genetic material, the field of molecular genetics has

been characterised by rapid advances during the last 65 years. With the sequencing of several genomes, including that of man, contemporary molecular genetics is now focussing on genotype – phenotype correlations as a means of identifying functions for each of the human genes.

C.1. The chemistry of life

Proteins are complex molecules, DNA is a monotonous polymer.

With the invention and development of analytical methodology such as paper chromatography, crystallography and spectroscopy, the stage was set to investigate the chemical nature of proteins and nucleic acids.

The notion of proteins as fundamentally important biological molecules dates back from the 1830s. It had been shown that there exist different proteins in nature, but that, upon hydrolysis, they all yield a class of simpler compounds, termed amino acids. All amino acids were isolated between 1819 ('oxide-caséeux'; that is, leucine)[1358,1359] and 1935 (threonine)[1360,1361], and in 1902, Emil Fischer (1852 – 1919) outlined the idea of the peptide bond, linking one amino acid with the next[1362]. In 1868, the Swiss biochemist Friedrich Miescher (1844 – 1895) discovered an unknown, phosphorus-rich acid substance in the nucleus, which he called nuclein (now known as nucleic acids)[1363]. The German medical doctor A. Kossel (1853 – 1927) identified the components of the nucleic acids: four nitrogenous bases, a five-carbon sugar and phosphoric acid. The differentiation into DNA and RNA dates back to the 1920s, when Phoebus A. Levene (1869 – 1940, Fig. X-1f) and W. Jones (1865 – 1935) showed that RNA contains ribose as a sugar[1364] and DNA contains deoxyribose[1365]. Moreover, RNA contains uracil as one of its bases, whereas DNA contains thymine (5-methyluracil). By the 1930s, it had been shown how the base, the sugar and the acid formed a nucleotide, and how several nucleotides were linked through phosphate diester bonds between their sugars to form polymers. Levene's finding in the 1920s that DNA contains approximately equal molar proportions of the four bases resulted in the tetranucleotide theory: the monomers in DNA consisted of a cyclic structure of the four bases, the sugars and the phosphate groups. Although the tetranucleotide theory was later interpreted in different ways, until the mid-1940s, DNA was essentially regarded as a monotonous polymer; such was Levene's influence.

C.2. The molecular nature of heredity

O. Avery: DNA is the carrier of hereditary information.

The observation that a harmless *S. pneumoniae* strain could be transformed to the virulent form by the addition of a cell-free extract of virulent bacteria led investigators to identify the substance responsible for this transformation. In 1944, Oswald T. Avery (1877 – 1955, Fig. X-1g) and co-workers met with stiff resistance in the scientific community when they published that DNA was the transforming substance and, hence, the carrier of hereditary information[1366]; DNA was still considered a simple, repetitive polymer that could not possibly carry any information. Instead, proteins were believed to be the carriers of genetic information, as they were thought to have the necessary complexity. It was argued that undetectable amounts of protein remaining in the DNA preparations of Avery were responsible for the transformation. The gradual acceptance that DNA indeed is the vehicle carrying hereditary information was brought about by the following observations:

- In 1950, Erwin Chargaff (1905 – 2002) showed that the four nucleotide bases are not necessarily present in DNA in equal amounts, as Levene had stated 30 years earlier. He also found that DNA's composition depends on the organism from which it has been extracted. Moreover, he mentioned that the molar ratios of adenine to thymidine and of guanine to cytosine were close to one[1367].
- Simultaneously with Chargaff's equivalence rule, as it became known, Alfred E. Mirsky (1900 – 1974) noticed that DNA is present in a constant amount in different somatic cells from an organism, and in half this amount in their germ cells. He pointed out that this ought to be the case if DNA is the genetic material[1368].
- Meanwhile, Avery showed that the transformation capacity of his preparation remained intact when treated with a variety of protein-degrading enzymes, but that this activity was completely lost when the preparation was treated very briefly with DNase, an enzyme hydrolysing DNA[1369].
- In 1952, Alfred D. Hersey (1908 – 1997) and Martha Chase (1928 – 2003) infected several *E. coli* cultures with ^{32}P (DNA)- or ^{35}S (protein)-labelled bacteriophages and showed that plaque-forming bacteria harboured the ^{32}P, but not the ^{35}S radioactivity. Thus, they showed that the phages' capability to reproduce within the host cells was dependent on the DNA, not on the protein[1370].

C.3. The structural basis of heredity

J. Watson and F. Crick: DNA is an anti-parallel double helix.

In the wake of the development of quantum mechanics, with its deep understanding of the chemical bond, in particular the nature of the hydrogen bond, and disillusioned with the application of physics during World War II, many physicists turned their attention towards biology in an attempt to unravel the basis of heredity. It was also in this mindset that Erwin Schrödinger (1887 – 1961), famous for his contributions to quantum mechanics, wrote his influential 'What is life?'[1371]. Against this background, James Watson (°1928, Fig. X-1h) and Francis Crick (1916 – 2004, Fig. X-1h) solved the structure of DNA[1372]. Their model was based on four pillars:

- From the shape of the molecules, the mathematician John Griffith (1928 – 1972) worked out that adenine and thymidine could fit together, linking up through a pair of hydrogen bonds, and that guanine and cytosine could pair up through three hydrogen bonds, both combinations resulting in equal molecular distances[1373].
- Chargaff's equivalence rule stated that [A] = [T] and [C] = [G].
- The symmetry of Rosalind E. Franklin's (1920 – 1958) superb X-ray crystallographies[1374] and the known density of the DNA molecule implicated a regular double helix.
- During a lecture in 1948, Linus Pauling (1901 – 1994, Fig. X-1i) stated that 'the gene' would consist of two congruent templates with complementary structures, each to 'serve as the mould for the production of a replica of the other part, and the complex of the two complementary parts thus can serve as the mould for the production of duplicates of itself'[1375].

Since DNA is the carrier of genetic information, Watson and Crick reasoned that any arbitrary sequence of the four nucleotides must be possible within the regularity of the helix, as seen in Franklin's X-rays. Since purines and pyrimidines have different sizes, they concluded that a purine and a pyrimidine would always have to form pairs to maintain the invariable diameter of the DNA helix, an idea supported by Griffith's finding and Chargaff's rule. And, in line with Pauling's idea, the DNA molecule had to be anti-parallel. In April 1953, Watson and Crick published their model of the structure of DNA, an anti-parallel double helix, and concluded their paper with their now famous statement:

'It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material'[1376].

C.4. Human molecular genetics

A. Garrod: inborn errors of metabolism & L. Pauling and V. Ingram: gene mutation affects a protein's characteristics, resulting in human disease.

Typically, genetic research investigates aberrations among a WT population. In this respect, the induction of mutations in genetic model organisms has proven extremely useful. Since inducing mutations in the human population is ethically unacceptable, human (molecular) genetics studies the inheritance of genetic disorders to unravel human gene function.

As early as 1902, Archibald E. Garrod (1857 – 1936) concluded from a study of family pedigrees that alkaptonuria (OMIM 203500), an arthritic condition accompanied by the excretion of wine-coloured urine, is a hereditary disease[1377]. To describe his observation, Garrod coined the term 'inborn errors of metabolism', as he proposed in 1908 that alkaptonuric individuals are homozygous for a recessive gene and that possession of this gene hampers some enzymatically catalysed metabolic reaction, resulting in the excretion of a purple metabolite[1378]. Thirty-three years later, Beadle and Tatum published their one-gene-one-enzyme experiments in agreement with Garrod's hypothesis (see X.B.3).

Molecular evidence that genes direct the primary structure of polypeptide chains, thus exerting their influence on protein characteristics (and hence, function), came from studies on sickle cell anemia (OMIM 603903). Analysis of the incidence in family pedigrees had shown that a single recessive mutant gene in the human genome must cause the condition. In 1949, Linus Pauling and co-workers demonstrated that the presence of an abnormal Hemoglobin molecule in red blood cells causes sickle cell anemia. From this finding, they concluded that sickle cell anemia is a gene-controlled molecular disease[1379]. This was at the time that Frederick Sanger (°1918, Fig. X-1i) developed the chemistry that allowed him to establish the primary structure of both chains of Insulin[1380]. Vernon M. Ingram (1924 – 2006, Fig. X-1i) applied the same strategy to sickle cell and normal Hemoglobin preparations, and in 1956, he announced that the mutant Hemoglobin carries a valine residue instead of glutamic acid at the sixth position from the amino terminus of the β-polypeptide chain[660]. This was the first proof that mutation in a gene can affect a protein's primary sequence and its characteristics, resulting in human disease. Ingram's work may be considered as the starting point of modern human molecular genetics, and its quest to understand the function of all human genes and their gene products.